

# Expression And Clinical Significance of BIRC5 In Low-Grade Gliomas Based On Bioinformatics Analysis

Xitong Yang (✉ [836826492@qq.com](mailto:836826492@qq.com))

The First Affiliated hospital of Dali University

---

## Primary research

**Keywords:** Low grade glioma, BIRC5, Tumor mutation burden, TCGA, Immunotherapy, Survival rate.

**Posted Date:** October 1st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-940151/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** To screen target genes analyze the expression and mechanism of target genes in Low-grade Gliomas (LGG).

**Methods:** LGG data were downloaded from TCGA database. Differentially expressed genes (DEGs) were screened by differential expression screening, protein–protein interaction (PPI) and enrichment analysis, and then the differentially immune related genes were obtained by analyzing DEGs and immune genes. The abundance of 22 leukocyte subtypes was calculated by CIBERSORT algorithm. The Hub gene constructed by PPI and the gene of prognosis model were analyzed to screen the LGG target gene. The correlation between target gene and LGG prognosis and clinical characteristics of LGG were analyzed.

**Results:** After immunotherapy, the survival of high TMB group was significantly longer than that low TMB group, and the scores of TMB group were correlated with the age and grade of patients. PPI indicates that AURKA, BUB1, AURKB, CDC20, BIRC5, TTK, CCNB2, CDCA8, and CENPE are the hub genes of LGG. The expression level of BIRC5 in LGG tumor tissues was significantly higher than that in normal tissues. The expression level of BIRC5 was correlated with the age and grade of LGG patients. BIRC5 might participate in the occurrence and development of LGG through P53 pathway.

**Conclusion:** BIRC5 can be used as LGG target gene, and its high expression in LGG indicates poor prognosis. BIRC5, as a therapeutic target of LGG, provides a new direction for immunotherapy of LGG.

## Introduction

Glioma, originating from neural stromal cells, is a frequent primary intracranial malignant tumor, accounting for 80% of central nervous system tumors and 40-50% of all malignant tumors in the brain [1,2]. According to classification of WHO, glioma can be classified as glioblastomas (GBMs) (WHO grade IV), lower-grade gliomas (LGGs) (WHO grade II and III) and so on. Low-grade glioma (LGG) refers to the second and third grade gliomas with isocitrate dehydrogenase (IDH) mutation in WHO classification, which has a good prognosis and mainly involves young patients [3]. LGG has the characteristics of slow growth, at present the treatment of LGG is still based on surgery, chemotherapy and radiotherapy are supplementary [4,5]. Despite adopting the best treatment measures, recurrent events remain inevitable [6]. In recent years, research on molecular biomarkers has brought new hopes and directions for accurate diagnosis, individualized treatment and prognosis of LGG. Therefore, there is an urgent need to find reliable and effective target genes as LGG therapeutic targets, which is of great significance for basic research and clinical treatment.

In the current study, we intends to download the transcriptome data and corresponding clinical data of low-grade gliomas from The Cancer Genome Atlas (TCGA) database, screen out LGG target genes through analysis of mutation burden and immune cell infiltration, use the clinical data for survival analysis and a series of comprehensive bioinformatics analysis, such as GSEA, explain the expression

and significance of target genes in LGG, so as to provide more theoretical basis for tumor biomarkers and potential therapeutic targets.

## Methods

### ***1. Data collection***

Clinical data, clinical data and mutation data were downloaded from the TCGA database (<https://portal.gdc.cancer.gov/>) then the data of three groups were sorted out. Transcriptome data was in the format of HTSeq-FPKM. There were 529 tumor tissues without normal samples. The clinical data included age, gender and grade of tumor patients. The mutation data of patients were selected as simple nucleus variation category and processed by Varscan software. Somatic mutation data were stored in the form of Mutation Annotation Format and analyzed using maftools [7]. The "maftools" software package was used for visualization, displaying the mutant genes in tumor samples, and performing characteristic analysis on the downloaded mutation data. This study was approved by the Ethics Committee of the First Affiliated Hospital of Dali University. The procedures followed were in accordance with the ethical standards of the Helsinki Declaration. Additionally, our research followed TCGA access rules and publication guidelines.

### ***2. Evaluation of TMB and its relationship with prognosis***

The tumor mutational burden (TMB) was estimated as the number of mutations per million bases, and the samples were divided into high and low TMB groups. We therefore defined TMB quantile > 80% as the high TMB group, and quantile  $\leq$  80% as the low TMB group [8,9]. The Perl script (java8 platform) was used for grouping. The Kaplan Meier statistical method was used to analyze the survival difference between the high and low TMB groups, and the Wilcoxon rank sum test was used to analyze the correlation between TMB and clinic.

### ***3. Identification of differentially expressed genes (DEGs), gene set enrichment analysis (GSEA) and PPI network construction.***

Differential expression genes between tumor tissues and normal tissues obtained from TCGA online database were identified by edgeR packages in R software. Adjusted p value  $\leq$  0.05 and  $|\log FC| \geq 1$  was set as the threshold for screening DEGs. In order to further study the correlation between immune genes and differentially expressed genes, the intersection between the above-mentioned differential genes (DEGs) and immune-related genes was selected, then the intersection was defined as the differential immune-related gene.

Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) were used to analyze the differentially expressed genes of TMB, and the "limma" package in r language software was used to predict the differentially expressed genes of TMB. The criteria for the statistically significant difference of DEGs was  $|\log FC| > 1$  in expression and FDR < 0.05. The heat map was constructed using the R package

"pheatmap". Seven software packages, including R software "colorspace", "string", "ggplot2", "dose", "org.hs.eg.db", "clusterprofiler" and "enrichment plot", were used to generate bar charts and volcanic maps, and the Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment were performed. Gene set enrichment analysis (GSEA) software was obtained from the website (<http://software.broadinstitute.org/gsea/index.jsp>) and worked on the JAVA platform. Using TMB data as the phenotype, "c2.cp.kegg.v7.4.symbols". "gmt gene set" (<http://software.broadinstitute.org/gsea/msigdb/>) was downloaded. STRING database (<http://string-db.org/cgi/input.pl>) is an online tool for evaluating the interaction information of protein. PPI network of DEGS was built by String V10 online tool (<https://string-db.org/>). By Cytohubba software of Cytoscape, the PPI network topology was calculated. According to the centrality score, the key points in the PPI network were identified, and the key pathogenic genes were derived.

#### ***4. Evaluation of difference of tumor infiltrating immune cells between high-risk group and low-risk group.***

CIBERSORT algorithm was used to estimate the proportion of 22 kinds of infiltrating immune cells. CIBERSORT can accurately estimate the immune components of tumor biopsy, and estimate the proportion of 22 human tumors infiltrating immune cells in samples by using gene expression data and pre-immune feature matrix [10]. CIBERSORT was run with 1,000 permutations and threshold < 0.05 as recommended. The expression differences of various immune cells in high TMB and low TMB groups were detected by R software "vioplot" package.

#### ***5. Construction of immune-related prognosis model.***

The characteristic genes were analyzed by univariate and multivariate Cox proportional hazard regression, and the prognosis prediction model was established. To verify the prognostic ability of the model, the individualized risk score was calculated by coefficient, and the patients were divided into high-risk group and low-risk group according to the median cut-off value of the risk score. Risk score formula was as follows: Risk score = expression of gene 1 \* coefficient + expression of gene 2 \* coefficient + ... expression of gene n \* coefficient [11]. R package "survival" was used to study the best critical value of risk score, and Kaplan-Meier survival curve was drawn. Kaplan-Meier survival curve showed the difference of survival time between high-risk group and low-risk group. The area under the curve was calculated by R package "survival ROC" to predict the survival time of the prognostic model and evaluate the accuracy and efficiency of the prognostic model [12]. In the survival analysis, the difference of P < 0.05 was statistically significant.

#### ***6. TIMER database analysis.***

The related genes were obtained by constructing prognosis model. The related genes were analyzed with different tumor infiltrating immune cells by SCNA in Timer2.0 ([http:// timer.comp-genomics.org](http://timer.comp-genomics.org)) database, and the correlation between immune cell content and gene copy number was analyzed. According to the expression level of genes related to prognosis model, patients and TCGA-HNSC data set were divided into high expression group and low expression group. Log-rank was used to test the

significance of the difference, and  $P < 0.05$  showed that the difference was statistically significant. In addition, the relationship between somatic copy number changes (SCNA) of these genes and tumor immune infiltration was evaluated by SCNA module in TIMER database. Bilateral Wilcoxon rank sum test was used to compare the infiltration level of each SCNA category with the normal level, and the difference was statistically significant ( $P < 0.05$ ).

### ***7. Correlation between target gene and immune checkpoint.***

The LGG target genes were screened out through PPI network construction and prognosis construction model, RNAseq SEQ data in LGG HTSeq-FPKM format was downloaded from TCGA database for log2 transformation, and the correlation between LGG target genes and CD274, PDCD1 and CTLA4 immune checkpoints was analyzed by using R software "ggplot2" package.

### ***8. Analysis of the correlation between target gene and LGG and prognosis.***

The expression data of LGG mRNA is downloaded from TCGA database, and the format was set as HTSeq-FPKM. The limma program package in R was used to extract target gene expression data from all genes, and the R software "ggplot2" program package was used to draw a histogram of target gene expression differences. LGG patients were divided into survival group and death group according to their survival status, and Kaplan-Meier statistical method was used to analyze the expression of target genes in the two groups. According to the median value of BIRC5 mRNA expression, the samples were divided into high expression group and low expression group. Perl script (JAVA8 platform) was used for grouping, and Wilcoxon rank sum test was used to analyze the correlation between target genes and clinical.

### ***9. GSEA analysis of target genes***

The expression data set. get file and phenotype data. els file of target gene were obtained by Perl language, and imported into GSEA4.0.3 software to analyze the relationship between target gene and GSEA. The GSEA software was obtained from the website (<http://software.broadinstitute.org/gsea/index.jsp>) and worked on the JAVA platform.

### ***10. Statistical analysis***

All statistical analyses were performed using R software (version 3.6.3). Kaplan-Meier method and cox regression analysis were used for survival analysis. Log-Rank test was used for curve comparison between groups. Statistical differences at  $P \leq 0.05$  were considered as significant.

## **Results**

### ***3.1 Analysis of lung cancer mutation database in TCGA database and characteristics of LGG patients.***

The single nucleotide variation data of LGG patients were obtained from TCGA database, the data processed by varscan software were selected, and the R software "maftools" package was used for

visualization. A total of 488 (96.44%) of the 506 LGG patients had mutations, of which missense mutations accounted for the largest proportion. The mutation types were dominated by single nucleotide polymorphisms. In the SNV classification, C>T was the most important type, and IDH1, TP53 and ATRX had the highest mutation rates, as shown in Fig 1A. The waterfall chart (Fig 1B) showed the specific mutation type and mutation gene ratio in each sample. Interaction diagram (Fig 1C) showed the relationship between two mutant genes, in which green indicated positive correlation and brown indicated negative correlation.

The clinical information of 515 LGG samples was obtained from TCGA database. After sorting out the data, it was found that there were 285 male patients and 230 female patients, and the clinical baseline was detailed in Table 1.

Table 1 Clinical characteristics of 515 LGG patients in TCGA clinical database.

<b>Variables</b>	<b>Number(%)</b>
<b>Vital status</b>	
Alive	406(78.83%)
Dead	109(21.17%)
<b>Age</b>	
≤65	483(93.79%)
>65	32(6.21%)
<b>Gender</b>	
Female	230(44.66%)
Male	285(55.34%)
<b>Tumor grade</b>	
WHO II	249(48.35%)
WHO III	265(51.46%)
Unknown	1(0.19%)

### ***3.2 Correlation between TMB score and survival outcome, clinical stage and primary tumor size.***

529 LGG transcriptome data downloaded from TCGA were divided into TMB groups, including 243 cases in high TMB group, 276 cases in low TMB group, and 10 cases were unable to match the grouping due to lack of data. It was found that the survival rate of high TMB group was significantly higher than that of low TMB group, and the difference was statistically significant ( $P<0.001$ ), which confirmed that the

survival time of patients in high TMB group was longer. According to the analysis of clinical data, the correlation between TMB scores and age, gender and grade were explored respectively. The results showed that TMB value of LGG patients aged > 65 years was significantly higher than that of LGG patients aged  $\leq 65$  years, and TMB value of LGG patients with G3 grade was significantly higher than that of G1-2 grade, the difference was statistically significant ( $P < 0.001$ ).

### ***3.3 Gene expression differential analysis, GSEA analysis and PPI network of different TMB groups.***

There were 318 differentially expressed genes between the high and low TMB groups, and the differentially expressed genes were identified by "limma" package, with  $|\log_2 FC| > 1$  and  $FDR < 0.05$  as screening criteria, and the top 10 differentially expressed genes were shown in Table 2. According to the differential gene FDR and its expression multiple, the volcano map was drawn by "ggplot2" package in Fig 3A. The heat map showed the expression level of DEGs by constructing the "pheatmap" package of R software, and the results were shown in Fig 3B. Venny diagram showed 1811 immune related genes and 318 differentially expressed genes, and 26 differentially expressed immune genes were obtained, as shown in Fig 3C. GO enrichment and KEGG pathway analysis of DEGs showed that BP was mainly related to pattern specification process, chromosome segregation, nuclear division and other pathways, while CC was mainly related to spindle, chromosome. Moreover, the MF was mainly related to DNA binding transcription activator activity-RNA polymerase II specific, microtubule binding and extracellular matrix structural constitution. KEGG pathway was related to cell cycle, proteoglycans in cancer, oocyte meiosis and other signal pathways. The results were shown in Fig 4A, B. In addition, GSEA results showed that the expression of adipocytokine signaling pathway, alanin aspartate and glutamate metabolism, aldosterone regulated sodium reabsorption and other signaling pathways were significantly enriched when TMB level was used as phenotype, as shown in Fig 4C.

Table 2 The top 10 DEGs in LGG

Gene	Low group	High group	LogFC	p-Value	FDR
SFRP2	77.20208	30.61399	-1.33445	5.86E-22	1.08E-17
SELL	47.09799	22.71912	-1.05176	2.61E-12	3.77E-10
USH1C	13.91831	6.485024	-1.1018	2.49E-14	1.07E-11
FSTL5	3.620375	1.527565	-1.24491	3.87E-20	2.38E-16
HPSE2	8.994663	3.922698	-1.19722	1.27E-16	1.67E-13
CHI3L1	27.82078	189.8005	2.770249	1.97E-11	1.81E-09
LTF	3.718816	24.61306	2.726509	4.72E-05	0.000281
FMOD	1.543689	11.3636	2.879966	7.54E-07	8.79E-06
AC004233.2	0.111767	5.472285	5.613576	1.50E-05	0.000108
OTP	0.05032	0.377971	2.909065	9.82E-18	1.88E-14

PPI network of DEGs was constructed by STRING v10, and visualized by Cytoscape. The result was shown in Fig 5A. The cytohubba software of Cytoscape was selected to determine the key nodes in the PPI network, and 10 key genes including CCNB1, AURKA, BUB1, AURKB, CDC20, BIRC5, TTK, CCNB2, CDCA8 and CENPE are obtained. The node degree score was generated according to Cytoscape, and the hub gene score was shown in Figure 5B.

### ***3.4 Immune cells in high and low TMB group.***

Based on the CIBERSORT algorithm, the proportion of 22 immune cells in each LGG patient is estimated, and the results are displayed in a block diagram, where different colors represent different cell subgroups, as shown in Fig 6A. The differences between the high and low TMB groups in 22 immune cells are illustrated in Fig 6B. Among them, the expression of the low mutation load group in Monocytes cells was significantly higher than that in the high mutation load group, and the difference was statistically significant. There was no significant difference in the expression of the two groups in the remaining cells.

### ***3.5 Establishment of risk model***

The related genes were obtained by univariate COX and multivariate COX analysis, a total of 9 genes, including CXCL11, GDF15, VEGFA, PLA2G2A, BIRC5, VAV3, IL9, TGFB2 and TNFRSF12A, were involved in the construction of the prognosis model. The survival time of the high expression group of 9 genes was significantly longer than that of the low expression group, and the difference was statistically significant ( $P < 0.05$ ). The results were shown in Fig 7A. The LGG patients were divided into high-risk and low-risk groups. Comparing the survival time of high-risk patients with that of low-risk patients, it was found that

the survival time of high-risk patients was significantly longer than that of low-risk patients, and the difference between the two groups was statistically significant ( $p < 0.05$ ), as shown in Fig 7B. The ROC curve constructed a model by calculating the area under the curve to predict the survival time of patients, and the AUC value of the curve was 0.863, indicating that the model had higher accuracy in predicting the survival time, as shown in Fig 7C.

### ***3.6 TIMER database analysis***

Immune cell analysis of high TMB and low TMB groups showed that the expression of low mutation load group was significantly higher than that of high mutation burden group in Monocytes cells, and 9 genes including CXCL11, GDF15, VEGFA, PLA2G2A, BIRC5, VAV3, IL9, TGFB2, TNFRSF12A participated in the model construction. The correlation between Monocytes cells and copy numbers of 9 genes was analyzed, and the results were shown in Fig 8A. Kruskal–Wallis analysis showed that among the 9 genes, the copy numbers of CXCL11, GDF15, PLA2G2A, TGFB2, VAV3 were correlated with the content of Monocytes cells.

Multivariate Cox regression analysis was performed on 26 immune-related DEGs by the "Survival" module of TIMER database. Nine genes were identified as highly correlated with the survival time of patients. The expression difference of these immune genes in LGG was verified by analyzing the expression of CXCL11, GDF15, VEGFA, PLA2G2A, BIRC5, VAV3, IL9, TGFB2 and TNFRSF12A in TIMER database. The results showed that SCNAs with different genes had influence on immune infiltration of LGG patients, as shown in Fig 8B. To illustrate the relationship between immune infiltrating cells and LGG survival rate, COX regression equation was used to calculate the expression levels of six immune infiltrating cells. Kaplan-Meier plots showed that the contents of 6 kinds of immune infiltrating cells were related to the survival and prognosis of patients, and the high levels of 6 kinds of cells could prolong the survival time of patients.

### ***3.7 The correlation between BIRC5 gene and immune checkpoint***

Ten Hub genes were constructed from PPI network and nine genes were constructed from survival model to screen BIRC5 as the target gene. 529 RNAseq data in LGG HTSeq-FPKM format were downloaded from TCGA database, and 1 case was deleted, with 528 cases in total. RNAseq data in FPKM format were transformed into  $\log_2$ , and the correlation between BIRC5 gene in LGG and CD274, PDCD1, CTLA4 immune checkpoints was analyzed by R software "ggplot2" package. The results indicated that BIRC5 gene expression was closely related to LGG prognosis, and BIRC5 gene expression was positively correlated with pdc1 and ctla4, but negatively correlated with CD274.

### ***3.8 The correlation between Birc5 expression and prognosis of LGG and clinical features of LGG.***

The expression data of BIRC5 single gene was extracted, and the columnar difference map of BIRC5 expression was drawn by software "ggplot2". The result was shown in Fig10A, which showed that the expression of BIRC5 in tumor tissue was significantly higher than that in normal tissue. The LGG patients

were divided into survival group and death group according to their survival status, and the correlation analysis was carried out between the two groups, as shown in Fig 10B. The results showed that the expression of BIRC5 in death group was significantly higher than that in survival group ( $p < 0.001$ ), indicating that BIRC5 gene was a risk factor for LGG patients, and the higher the expression of BIRC5, the worse the prognosis of LGG patients. According to the median expression of BIRC5 mRNA, 528 LGG samples were divided into high expression group and low expression group. The relationship between BIRC5 expression and clinical characteristics of LGG patients was expounded. The results showed that BIRC5 expression was related to age and grade of LGG patients. The expression of BIRC5 in patients aged  $> 40$  with grade G3 was significantly higher than that in patients aged  $\leq 40$  with G2 grade, and the difference was statistically significant, as shown in Fig 10C for details, which indicated that the expression of BIRC5 was correlated with advanced age and G3 stage.

### 3.9 GSEA enrichment analysis of BIRC5 gene

To further understand the biological process and function of BIRC5 in LGG and the related signal pathway, LGG patients were divided into high expression group and low expression group according to BIRC5 expression level. In this study, BIRC5 expression data set. get file and phenotypic data. els file were obtained by Perl language and imported into GSEA4.0.3 software. The result showed that the high expression of BIRC5 mRNA was related to P53 signaling pathway, DNA replication, cell cycle, phosphoinositol signaling system and neuroactive ligand receptor interaction signaling pathway. The results were shown in Fig 11 and Table 3.

Table 3 GSEA analysis of B1RC5 enrichment results

ID	ES	NES	p.adjust	FDR
CELL_CYCLE	0.797	2.771	0.013	0.009
P53_SIGNALING_PATHWAY	0.723	2.296	0.013	0.009
DNA_REPLICATION	0.807	2.306	0.013	0.009
PHOSPHATIDYLINOSITOL_SIGNALING_SYSTEM	-0.577	-2.008	0.013	0.009
NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION	-0.565	-2.349	0.016	0.011

## Discussion

Recently, the morbidity of gliomas has been rising annually, which is characterized by low cure rate, high incidence and recurrence rate [13]. Immunotherapy for glioma has been studied in clinic for a long time, but there is still a lack of reliable molecular biomarkers to judge the sensitivity of immunotherapy [14]. Therefore, the identification of prognostic biomarkers related to immunity and the screening of target

genes are particularly important in tumor immunotherapy, and the screening of target genes can be used as potential therapeutic targets.

TMB is the total number of mutations per megabase in the exon coding region of the evaluated gene in the sample [15]. Currently, it is generally believed that genome variation is the main cause of LGG [16]. In the present study, BIRC5 was considered as an effective biomarker and independent risk gene for LGG through TMB and immune cell infiltration analysis, which could be used as an immunotherapy target. The results of mutation load study showed that the survival time of high TMB was significantly better than that of low TMB group, and the difference was statistically significant. In addition, our study illustrated that age and grade were significantly correlated with TMB scores. LGG patients aged > 65 years with grade G3 had higher TMB value, and the survival time of high TMB group was longer, which indicated that the higher TMB value, the better the immune response effect of LGG patients. GO and KEGG functional enrichment analysis showed that pattern specification process, spindle, DNA binding transcription activator activity-RNA polymerase II - specific, and cell cycle signaling pathway were associated with LGG patients. GSEA analysis showed that more immune-related biological processes were enriched in high TMB group, which indicated that high TMB enhanced immune phenotype.

Cancer-related inflammatory reactions are closely related to the occurrence and development of tumors. In tumor microenvironment, changes in the number of immune cells related to inflammatory infiltration can cause oxidative damage and abnormal repair of DNA, resulting in changes at the gene level or epigenetic level, eventually leading to promote the growth, invasion and metastasis of glioma [17]. The immune results in this study showed that the expression of monocytes cells in low mutation burden group was significantly higher than that in high mutation burden group. Monocytes cells has been demonstrated to inhibit acquired immunity and promote angiogenesis and proliferation, invasion and metastasis of tumor cells [18]. Meanwhile, previous studies suggested that the survival time of high TMB group was better than that of low TMB group, while the expression of Monocytes cells in low TMB group was higher than that in high TMB group in immune cells, which was consistent with previous research results, namely, the low TMB group with monocytes cell expression promoted tumor metastasis and had shorter survival time. Univariate COX and multivariate COX were used to analyze the differential immune genes, and the prognosis model was constructed. The results showed that 9 genes participated in the model construction. The risk survival curve showed that the survival time of high-risk patients was longer than that of low-risk patients, and the difference was statistically significant. The ROC curve (AUC = 0.863) suggested that the prognosis model had higher accuracy. The expression levels of CXCL11, GDF15, VEGFA, PLA2G2A, BIRC5, VAV3, TGFB2, TNFRSF12A in the genes involved in model construction were positively correlated with survival time, while the expression level of IL9 was negatively correlated with survival time. The results of TIMER database analysis showed that 9 immune genes were related to LGG immune infiltration, while immune cells were related to LGG survival rate. A total of 9 genes could affect the survival rate of patients by being related to immune cells. Previous studies have shown that the expression of monocytes cells in immune cells was related to LGG tumor metastasis and survival time of patients. Therefore, the copy numbers of genes related to the construction of prognosis model in

monocytes cells were analyzed, and the results showed that the copy numbers of CXCL11, CDF15, PLA2G2A, TGFB2 and VAV3 genes were related to the content of monocytes cells.

10 Hub genes were obtained by constructing PPI network, and nine genes related to LGG were obtained by constructing prognosis model. After screening, it was found that BIRC5 gene expression was closely related to LGG diseases and prognosis, and could be used as a target gene for LGG treatment. Baculoviral inhibitor of apoptosis protein repeat containing protein 5 (BIRC5), a member of inhibitor of apoptosis protein (IAP) family, can be used as a negative regulatory protein to inhibit apoptotic cell death, regulate cell mitosis, inhibit caspase-3 and caspase-7 activities, so as to inhibit tumor cell apoptosis and promote cell proliferation [19,20]. BIRC5 is highly expressed in a variety of tumor tissues, while its expression level is very low or not expressed in normal tissues. Some studies have confirmed that BIRC5 is involved in the malignant transformation of tumor cells and the physiological process of anti-apoptosis, suggesting that BIRC5 gene can be used as an important target for anti-tumor therapy [21]. In this study, BIRC5 expression in LGG normal tissues and tumor tissues was compared and analyzed. BIRC5 expression in LGG tumor patients was significantly higher than that in normal tissues. BIRC5 expression in LGG patients was related to prognosis, and BIRC5 expression in death population was significantly higher than that in survival population, which was consistent with previous research results. Therefore, BIRC5 gene can be considered as an independent risk factor for LGG.

Immunocheckpoint regulators-cytotoxic t lymphocyte antigen-4 (CTLA4), programmed cell death protein 1 (PD-1), and programmed cell death ligand 1 (PD-L1), also known as CD274, is an important immune checkpoint molecules, which can be used as important target molecules for tumor immunotherapy. The results of immune checkpoint analysis indicated that BIRC5 show a positive correlation with PDCD1 and CTLA4, and a negative correlation with CD274. In LGG, it may be involved in the tumor immune escape process mediated by CD274/PDCD1/CTLA4 in the course of LGG disease. IAP, a highly conserved anti-apoptosis factor, mainly inhibits apoptosis by inhibiting caspase activity and regulating NF- $\kappa$ B. Apoptosis is of great significance for maintaining the normal function of cells. In normal condition, the body can clear damaged DNA or cells with abnormal circulation cycle by starting apoptosis program. However, once the cells become malignant and the apoptosis mechanism is damaged, the body cannot actively remove the malignant cells, eventually result in tumorigenesis [22]. As a member of the IAP family, BIRC5 has a significant effect on inhibiting cell apoptosis, which can directly or indirectly inhibit the activity of caspase-3 and caspase-7 through a variety of ways, and ultimately prevent cell apoptosis [21], promote angiogenesis. Therefore, BIRC5 can be used as a biomarker for tumor diagnosis, treatment and prognosis [23].

The correlation analysis between BIRC5 and the clinical characteristics of LGG patients showed that the expression of BIRC5 was significantly increased in patients aged > 40 years with grade G3. To further explain the possible molecular mechanism of BIRC5 in the occurrence and development of LGG, the present study analyzed the data obtained from public databases using GSEA to identify its important pathway. The results showed that the high expression of BIRC5 gene could inhibit P53 signaling pathway, DNA replication, cell cycle and signaling pathway. P53 gene mutation is the most common gene mutation

in glioma [24], which can inhibit cancer by regulating cell growth, maintaining DNA integrity and preventing gene mutation cells with canceration tendency, and is also the most studied tumor gene at present. Available studies indicate that the occurrence and development of various tumors are often accompanied by P53 gene mutation [25]. The key function of tumor suppressor gene P53 is to repair DNA damage. In addition, P53 initiates programmed cell death when DNA damage is serious and cannot be repaired. There is a close relationship between P53 and BIRC5. wild type P53 has a negative regulatory effect on BIRC5, and wild type P53 inhibits BIRC5 at mRNA and protein levels. Once P53 is mutated, the repressor protein complex cannot be formed and bound to the BIRC5 promoter, resulting in abnormal expression of BIRC5 [26], promoting the occurrence and development of tumors. In this study, the mechanism of the correlation between BIRC5 and LGG may be related to p53 pathway, and the specific mechanism needs to be further explored.

Taken together, high mutation load is positively correlated with the survival and prognosis of LGG patients. The patients were divided into low-risk group and high-risk group by constructing the prognosis model. The results show that the high-risk group were more likely to benefit from immunotherapy. By analyzing the results of TMB and immune cell infiltration, BIRC5 was chosen as an independent risk factor of LGG. The results showed that the high expression of BIRC5 in LGG patients had a worse prognosis. Therefore, the expression of BIRC5 can be used as one of the prognostic indicators of LGG patients. BIRC5 may participate in the occurrence and development of LGG tumors through P53 signaling pathway. Inhibition of BIRC5 expression may become a new direction of LGG targeted therapy in the future.

There were limitations in the present study. (1) This study is a retrospective study, and the correlation between TMB-related immune genes and immune cell infiltration needs to be confirmed by experiments. (2) There is a lack of many clinical samples to verify the prognostic effect of TMB and its potential relationship with immune infiltration; (3) The correlation between BIRC5 gene and LGG needs more research to confirm. Therefore, more studies are needed to further confirm the mechanism and expression significance of BIRC5 in LGG.

## Conclusion

This study found that mutation burden is related to the survival of LGG patients with immunotherapy. The higher the TMB value, the longer the survival time. The TMB value is closely related to the age and grade of LGG patients. BIRC5 may play a key role in the development of LGG through the P53 signaling pathway. The higher the expression level of BIRC5, the worse the prognosis of patients. Its expression level is related to tumor grade and age. BIRC5 expression level can be used as one of the prognostic indicators for judging LGG patients. This study provides more basis and ideas for the mechanism of BIRC5 in LGG through bioinformatics research, and provides new ideas for LGG targeted therapy.

## Declarations

**Ethics approval and consent to participate:** The research was under the approval of the Ethics Committee of the First Affiliated Hospital of Dali University.

**Consent for publication:** Consent.

**Availability of data and materials:** The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Competing interests:** The authors declare that they have no competing interests.

**Funding:** This study was supported by the National Natural Science Foundation of China (No. 81360206); and Scientific research fund project of Yunnan education department (2019J0775) and Science and Technology Project of Dali City (2019KGB052), Colleges joint specific project in Yunnan Province (2019 FH001-(020)).

### Authors' contributions

XTY: project administration, data curation, manuscript writing.

XTY: methodology, investigation, manuscript writing.

XTY: formal analysis, methodology, software.

XTY conceptualization, supervision, manuscript reviewing and editing.

XTY read and approved the final manuscript.

Acknowledgements: None.

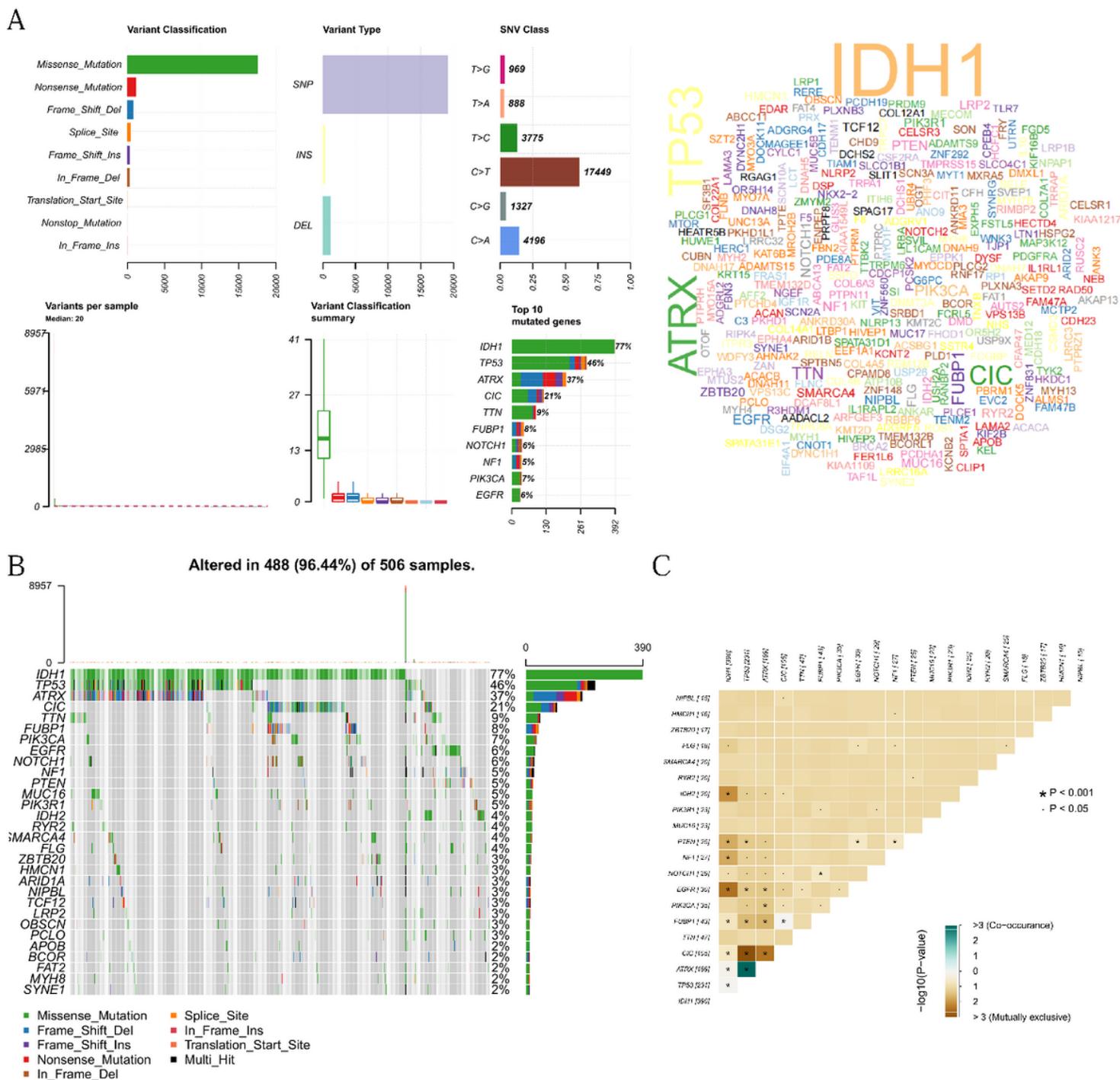
## References

1. Ostrom QT, Gittleman H, Truitt G, Boscia A, Kruchko C, Barnholtz-Sloan JS. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2011-2015. *Neuro Oncol.* (2018) 20 (suppl\_4): iv1-iv86. doi: 10.1093/neuonc/noy131.
2. Wen PY, Reardon DA. Neurooncology in 2015: progress in glioma diagnosis, classification and treatment. *Nat Rev Neurol.* 2016;12(2):69–70. doi:10.1038/nrneurol.2015.242.
3. Youssef G, Miller JJ. Lower grade gliomas. *Curr Neurol Neurosci Rep.* 2020;20(7):21. doi:10.1007/s11910-020-01040-8.
4. Acharya S, Liu JF, Tatevossian RG, Chiang J, Qaddoumi I, Gajjar A, et al. Risk stratification in pediatric low-grade glioma and glioneuronal tumor treated with radiation therapy: an integrated clinicopathologic and molecular analysis. *Neuro Oncol.* 2020;22(8):1203–13. doi:10.1093/neuonc/noaa031.

5. Darlix A, Rigau V, Fraisse J, Gozé C, Fabbro M, Duffau H. Postoperative follow-up for selected diffuse low-grade gliomas with WHO grade III/IV foci. *Neurology*. 2020;94(8):e830–41. doi:10.1212/WNL.0000000000008877.
6. Wang J, Cheng P, Pavlyukov MS, Yu H, Zhang Z, Kim SH, et al. Targeting NEK2 attenuates glioblastoma growth and radioresistance by destabilizing histone methyltransferase EZH2. *J Clin Invest*. 2020;130(9):5027. doi:10.1172/JCI142404.
7. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747–56. doi:10.1101/gr.239244.118.
8. Xu Z, Dai J, Wang D, Lu H, Dai H, Ye H, et al. Assessment of tumor mutation burden calculation from gene panel sequencing data. *Onco Targets Ther*. (2019) 12:3401-3409. doi: 10.2147/OTT.S196638. eCollection 2019.
9. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types[J]. *Nat Genet*. 2019;51(2):202–6. doi:10.1038/s41588-018-0312-8.
10. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–7. doi:10.1038/nmeth.3337.
11. Chen HY, Yu SL, Chen CH, Chang GC, Chen CY, Yuan A, et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med*. 2007;356(1):11–20. doi:10.1056/NEJMoa060096.
12. Lorent M, Giral M, Foucher Y. Net time-dependent ROC curves: a solution for evaluating the accuracy of a marker to predict disease-related mortality. *Stat Med*. 2014;33(14):2379–89. doi:10.1002/sim.6079.
13. Ishihara M, Yamamoto K, Miwa H, Nishi M. Spontaneous complete regression of a brain stem glioma pathologically diagnosed as a high-grade glioma. *Childs Nerv Syst*. 2017;33(12):2177–80. doi:10.1007/s00381-017-3570-3.
14. Kamran N, Alghamri MS, Nunez FJ, Shah D, Asad AS, Candolfi M. Current state and future prospects of immunotherapy. *Immunotherapy*. 2018;10(4):317–39. doi:10.2217/imt-2017-0122.
15. Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol Cancer Ther*. 2015;14(4):847–56. doi:10.1158/1535-7163.MCT-14-0983.
16. Zhang CM, Brat DJ. Genomic profiling of lower-grade gliomas uncovers cohesive disease groups: implications for diagnosis and treatment. *Chin J Cancer*. 2016;35:12. doi:10.1186/s40880-015-0071-1.
17. Sowers JL, Johnson KM, Conrad C, Patterson JT, Sowers LC. The role of inflammation in brain cancer. *Adv Exp Med Biol*. 2014;816:75–105. doi:10.1007/978-3-0348-0837-8\_4.
18. Condeelis J, Pollard JW. Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell*. 2006;124(2):263–6. doi:10.1016/j.cell.2006.01.007.

19. Cho M, Lee OH, Chang EM, Lee S, Moon S, Lee J, et al. BIRC5 expression is regulated in uterine epithelium during the estrous cycle. *Genes (Basel)*. (2020) 11(3):282. doi: 10.3390/genes11030282.
20. MacDonald JA, Kura N, Sussman C, Woods DC. Mitochondrial membrane depolarization enhances TRAIL-induced cell death in adult human granulosa tumor cells, KGN, through inhibition of BIRC5. *J Ovarian Res*. 2018;11(1):89. doi:10.1186/s13048-018-0463-3.
21. Shi J, Tan SY, Lee AZE, Zhang S, Sasidharan SL, Wong B, et al. Restoring apoptosis dysregulation using survivin inhibitor in nasopharyngeal cancer. *Head Neck*. 2020;42(5):913–23. doi:10.1002/hed.26068.
22. Ambrosini G, Adida C, Altieri DC. A novel anti-apoptosis gene, survivin, expressed in cancer and lymphoma. *Nat Med*. 1997;3(8):917–21. doi:10.1038/nm0897-917.
23. Cai C, Wang W, Tu Z. Aberrantly DNA methylated-differentially expressed genes and pathways in hepatocellular carcinoma. *J Cancer*. (2019) 10 (2):355-366. doi: 10.7150/jca.27832. eCollection 2019.
24. Rao S, Kanuri NN, Nimbalkar V, Arivazhagan A, Santosh V. High frequency of H3K27M immunopositivity in adult thalamic glioblastoma. *Neuropathology*. 2019;39(2):78–84. doi:10.1111/neup.12537.
25. Meder L, König K, Ozretić L, Schultheis AM, Ueckerth F, Ade CP, et al. NOTCH, ASCL1, p53 and RB alterations define an alternative pathway driving neuroendocrine and small cell lung carcinomas. *Int J Cancer*. 2016;138(4):927–38. doi:10.1002/ijc.29835.
26. Liang H, Huang J, Li H, et al. Expression levels of the apoptotic inhibition molecules BIRC5 in non-small cell lung cancer and its correlation with Tp53 expression. *Journal of Modern Oncology*. (2013) 21 (7): 1496–9. doi: 10.3969/j.issn.1672-4992.2013.07.23.

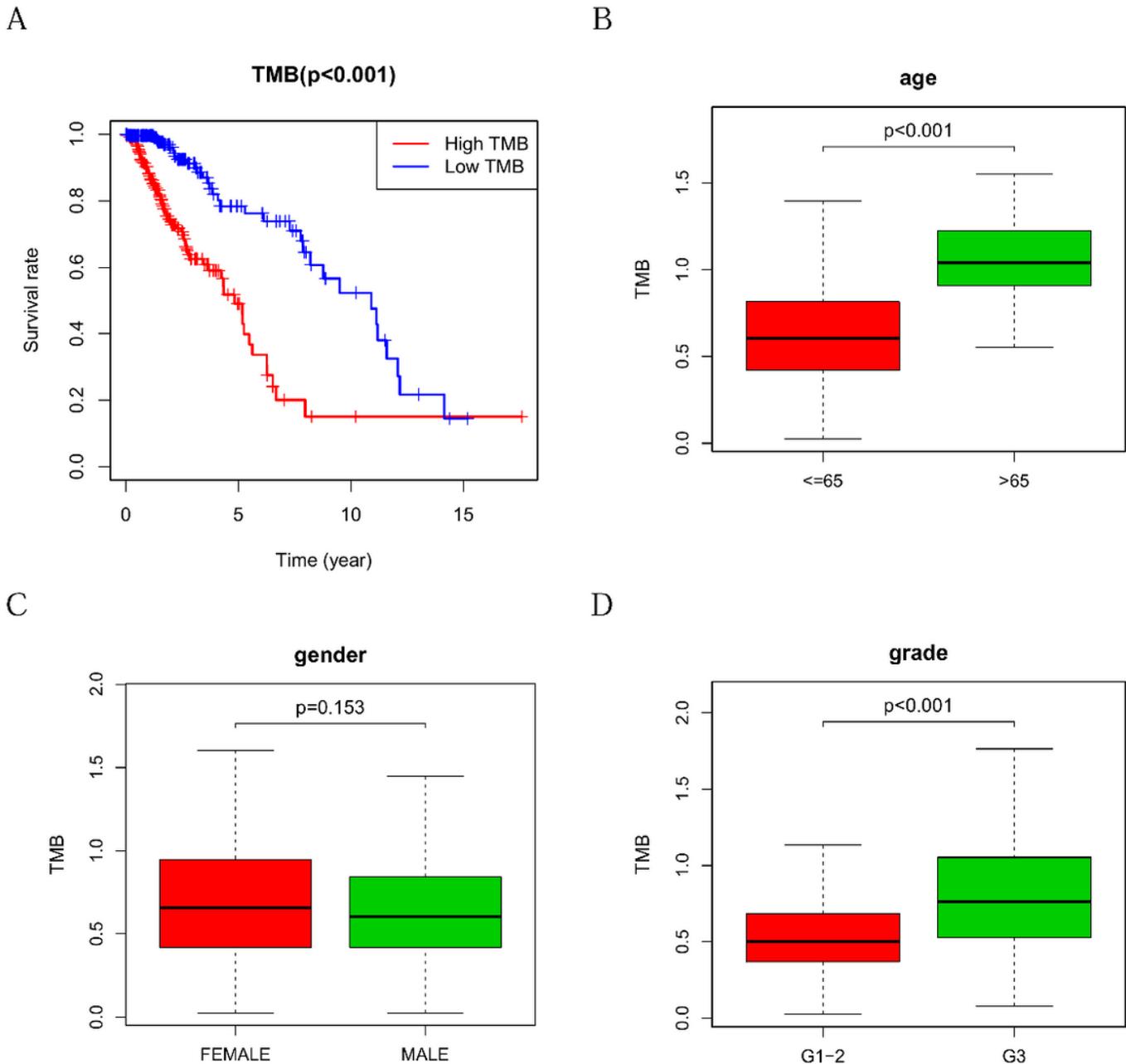
## Figures



**Figure 1**

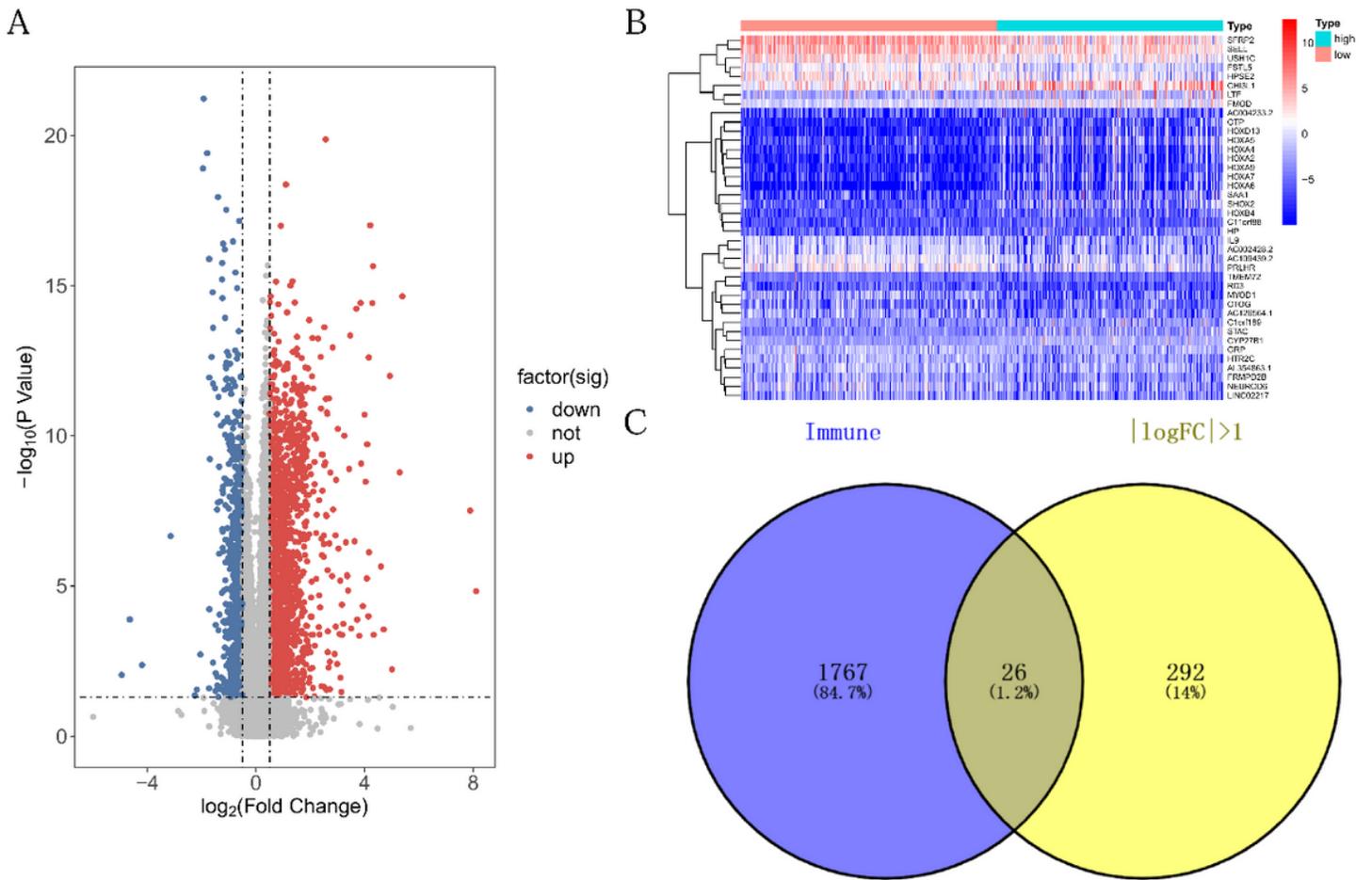
A. overview of LGG gene mutations: mutation classification; Classification of SNV; Variants per sample Variant Classification summary; Top 10 mutated genes; Gene cloud; B. waterfall chart. The mutation information of each gene in each sample was shown in the figure. The ordinate represented gene names, the abscissa represented different samples, the upper bar graph represented mutation load quantity, the following notes indicated mutation types, and the right bar graph indicated the distribution of mutation

types among the top 30 genes; C. Mutation correlation analysis. Exclusivity and correlation between mutations. Green represented Co-occurrence, and gray represented mutually exclusive.



**Figure 2**

Prognostic analysis of tumor mutation load and its correlation with clinical factors. A. correlation between high-low TMB groups and survival rate ( $p < 0.001$ ); B. Correlation between age group and TMB score; C. Correlation between gender group and TMB score; D. Correlation between grade stage and TMB score;



**Figure 3**

A. Volcano map of differential genes. Notes:  $|\log_2\text{FC}| > 0.5$ ,  $\text{FDR} \leq 0.05$ . Red represented a significant increase, while green represented a significant decrease; B. The top 40 genes (up-regulated and down-regulated genes) were shown in the heat map, and the genes were distributed in high and low TMB groups: Red represented higher expression level, blue represented lower expression level, green represented high TMB group and red represented low TMB group; C. Venny diagram of immune related genes and DEGs.

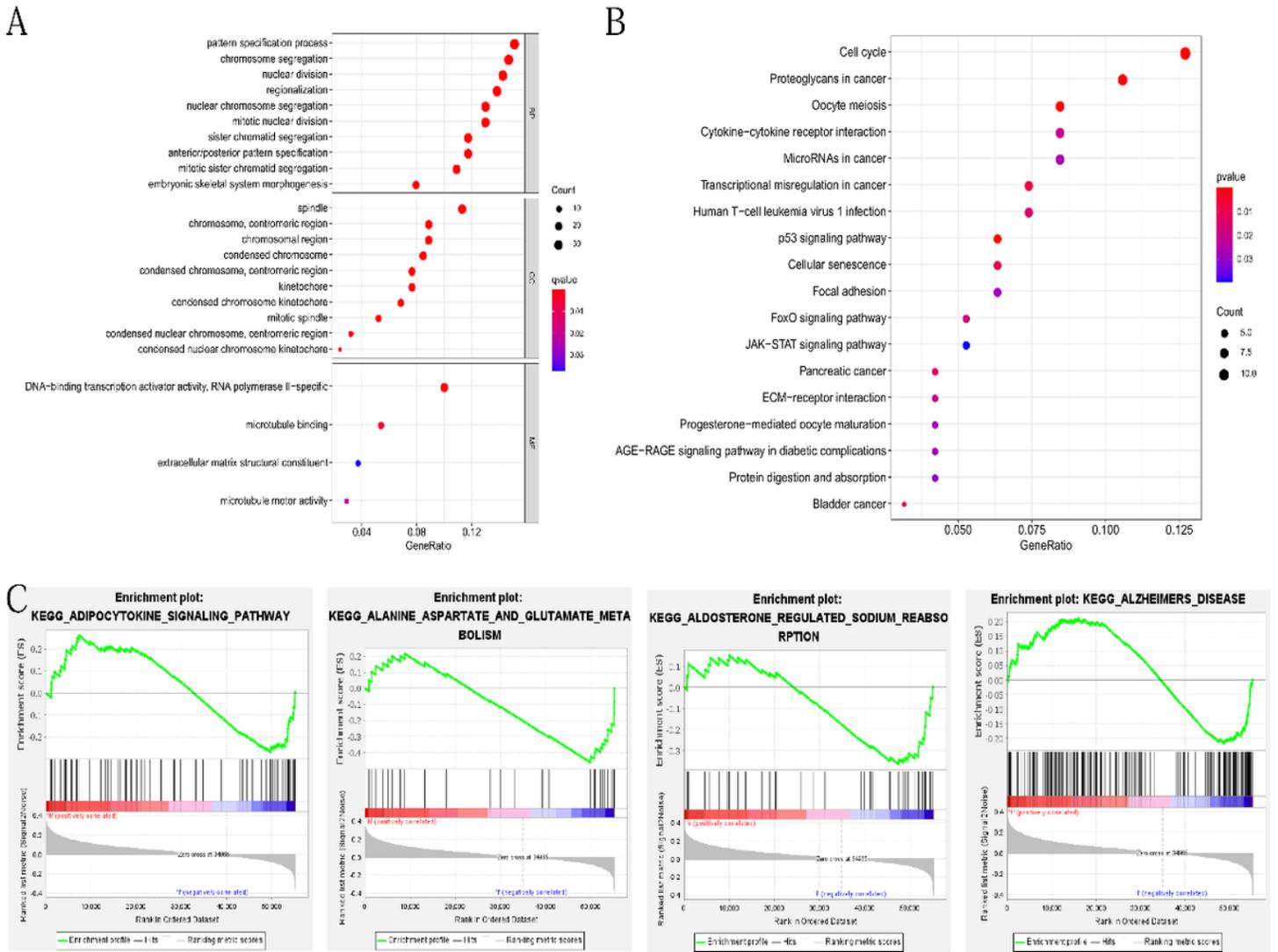
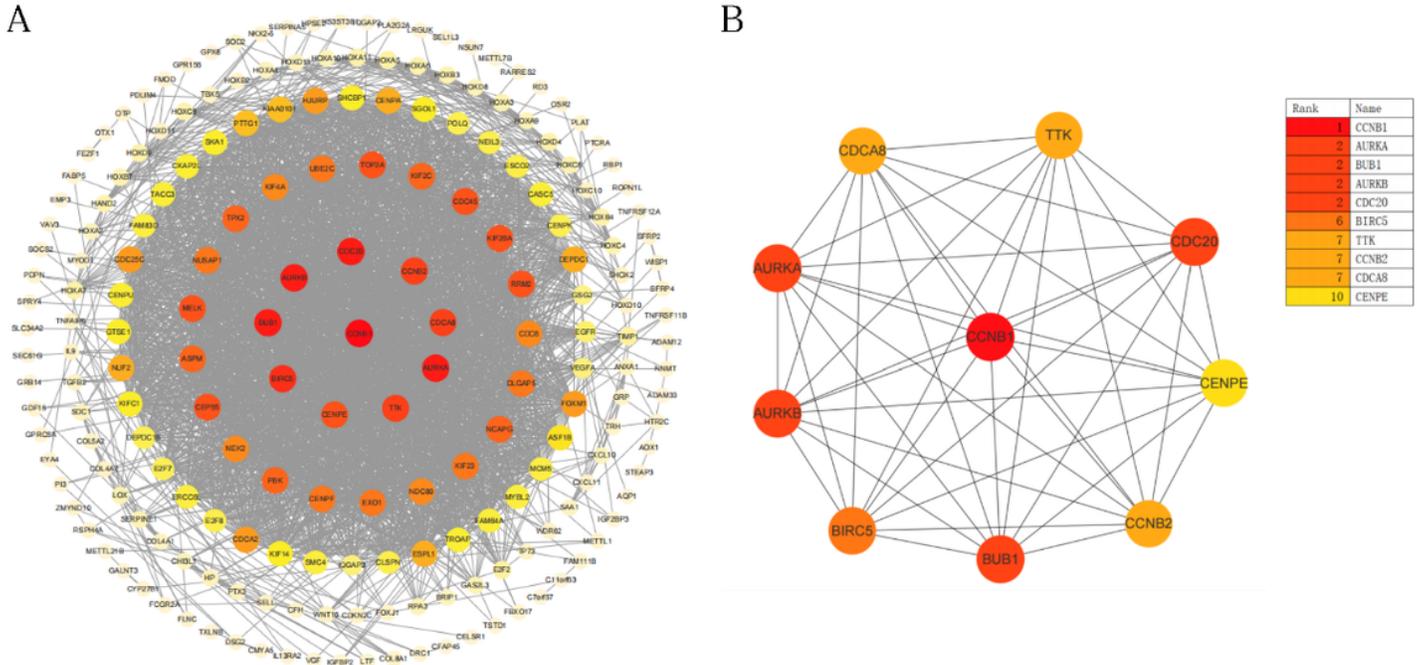


Figure 4

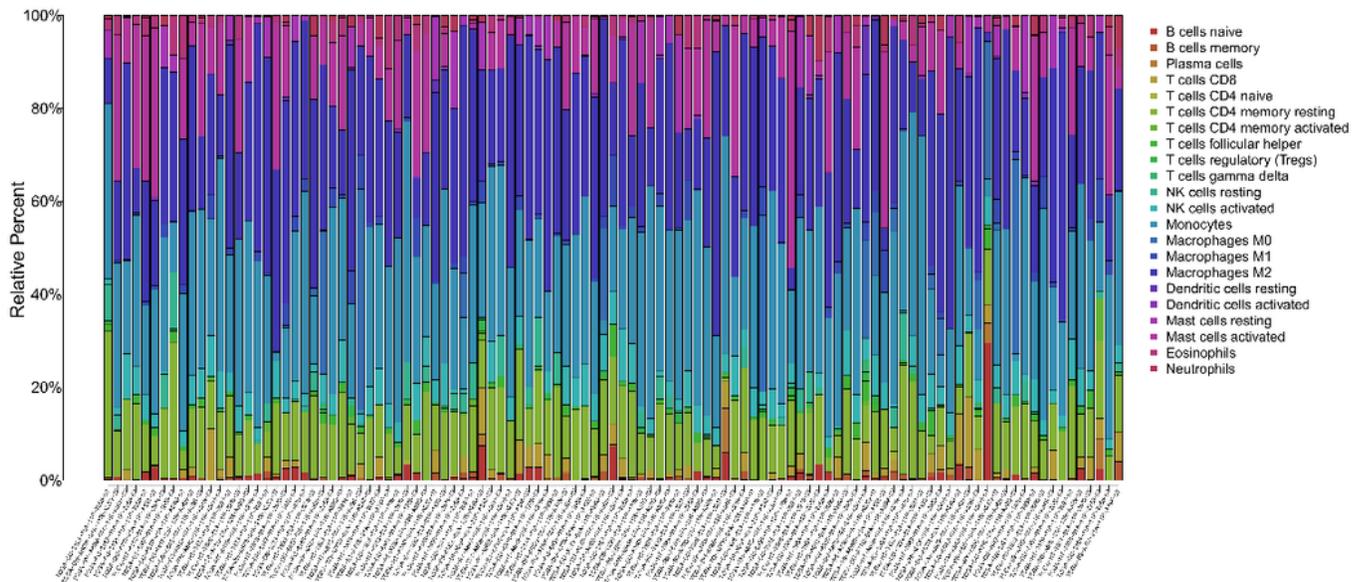
A. GO enrichment analysis; B. KEGG analysis results; C. GSEA analysis results. GSEA showed the top TMB-related signaling axis, FDR <0.4.



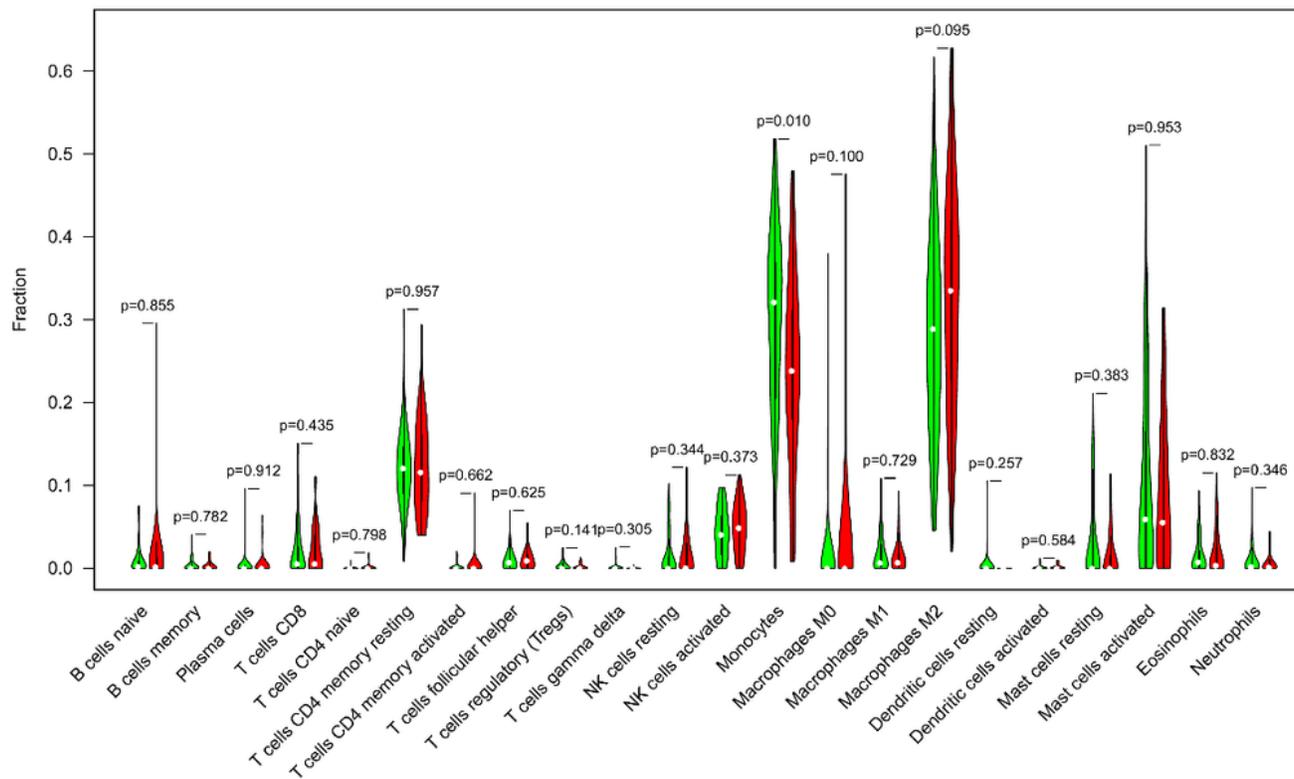
**Figure 5**

A. Protein–protein interaction network analyzed using Cytoscape software. The size and color of that correspond nodes of each gene were determined according to the degree of interaction. The closer to the red node, the higher the connectivity between two nodes; B. Protein–protein interaction network for the top 10 hub genes. Node color indicated the number of degrees. The top 10 ranked hub genes were depicted using a pseudocolor scale. Red represented highest degree, and yellow represented lowest degree.

A



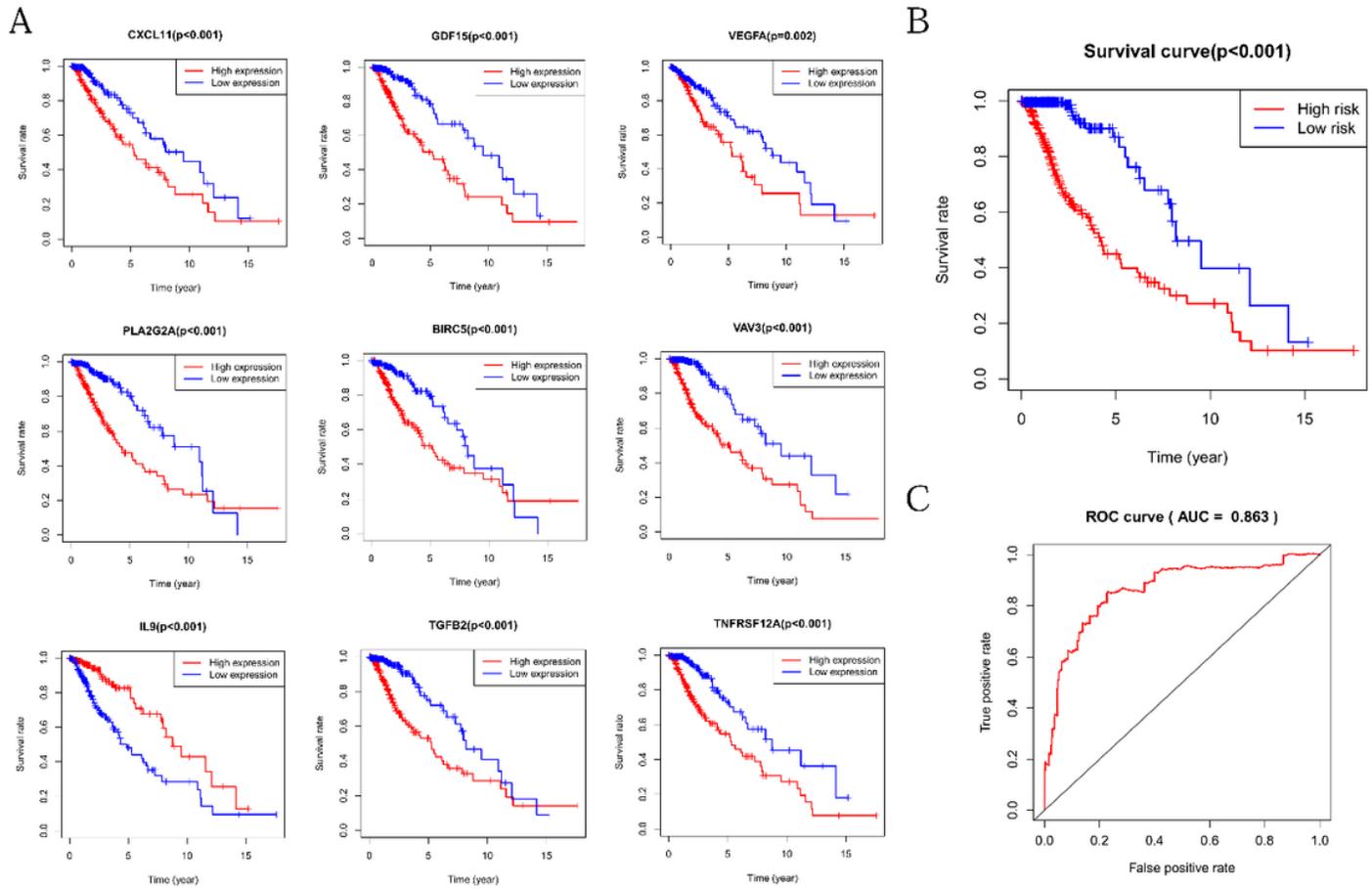
B



**Figure 6**

A. Bar graph of immune cell content. A summary of the estimated scores of 22 immune cell subtypes in the CIBERSORT algorithm. Each bar graph showed the proportion of each patient's cell, each color represented 22 kinds of immune cells, and there were notes below the legend; B. Violin diagram of immune cell infiltration, green represented low mutation load group, red represented high mutation load

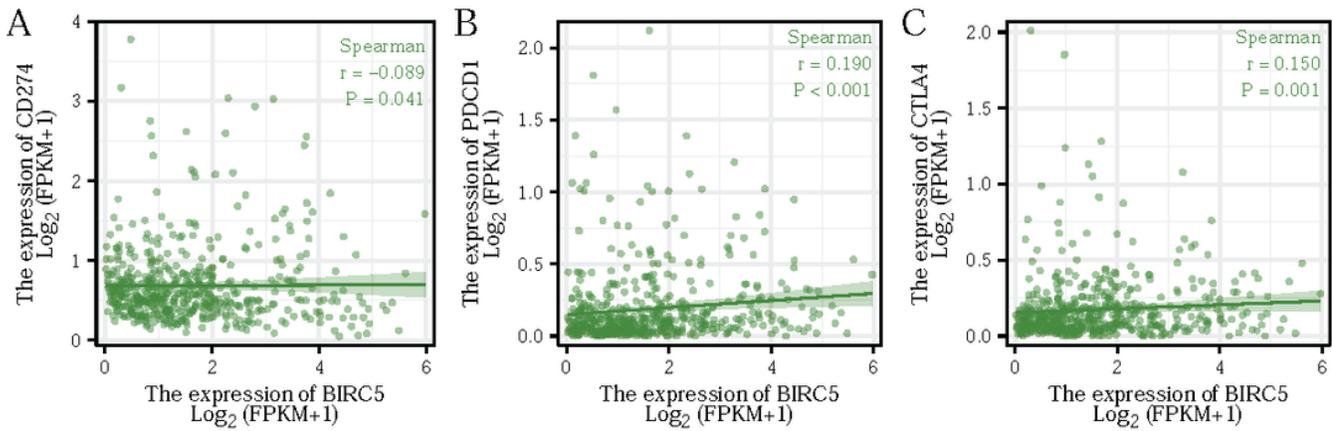
group,  $P < 0.05$  indicated that there was significant difference in immune cell expression between high and low mutation load groups.



**Figure 7**

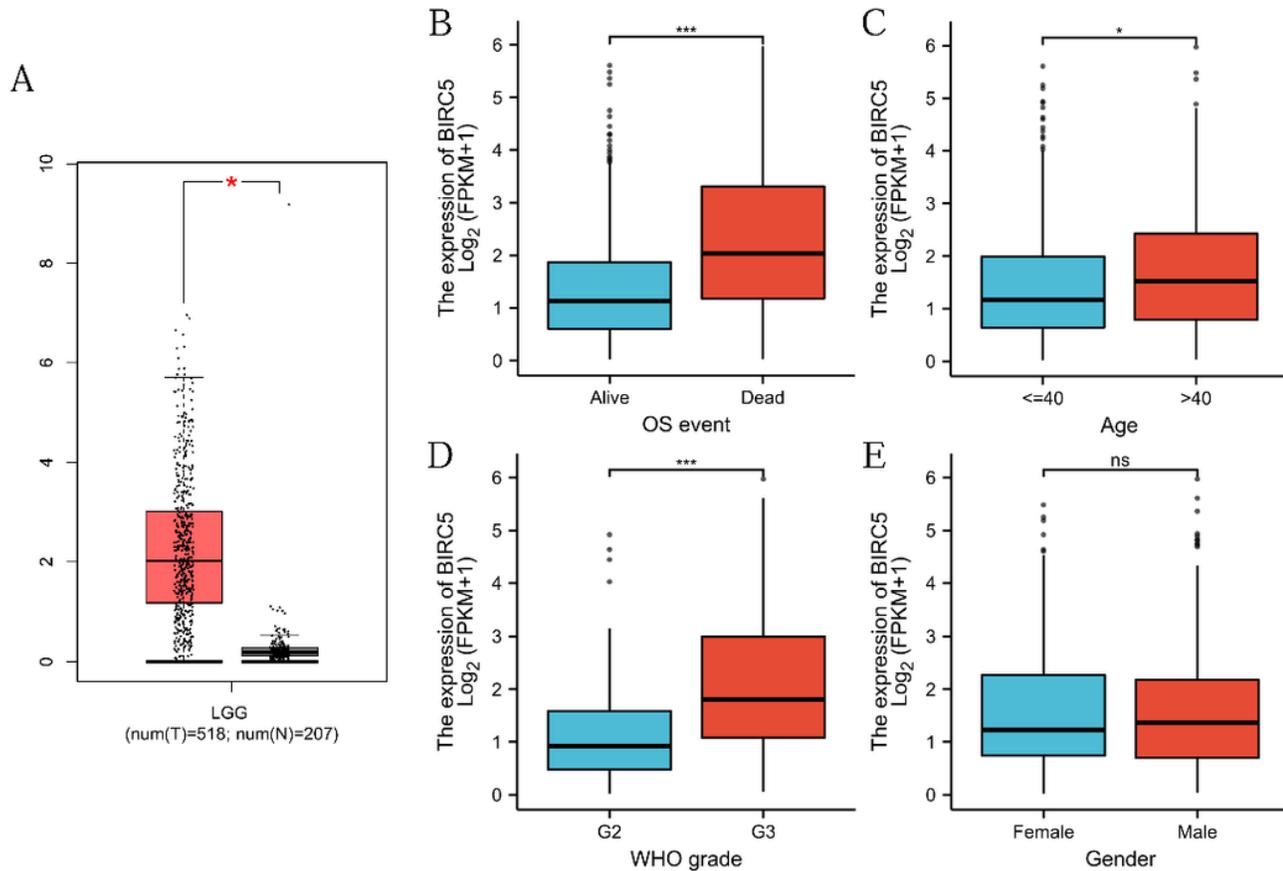
A: 9 Gene survival analysis of the constructed model; B. Risk score survival analysis; C. ROC curve model;





**Figure 9**

Analysis of the correlation between BIRC5 and CD274, PDCD1, CTLA4 expression



**Figure 10**

A. BIRC5 expression histogram difference between LGG tissue and normal tissue; B. BIRC5 expression difference analysis in LGG patients; C-E. BIRC5 expression was correlated with the clinical characteristics of LGG patients; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ;

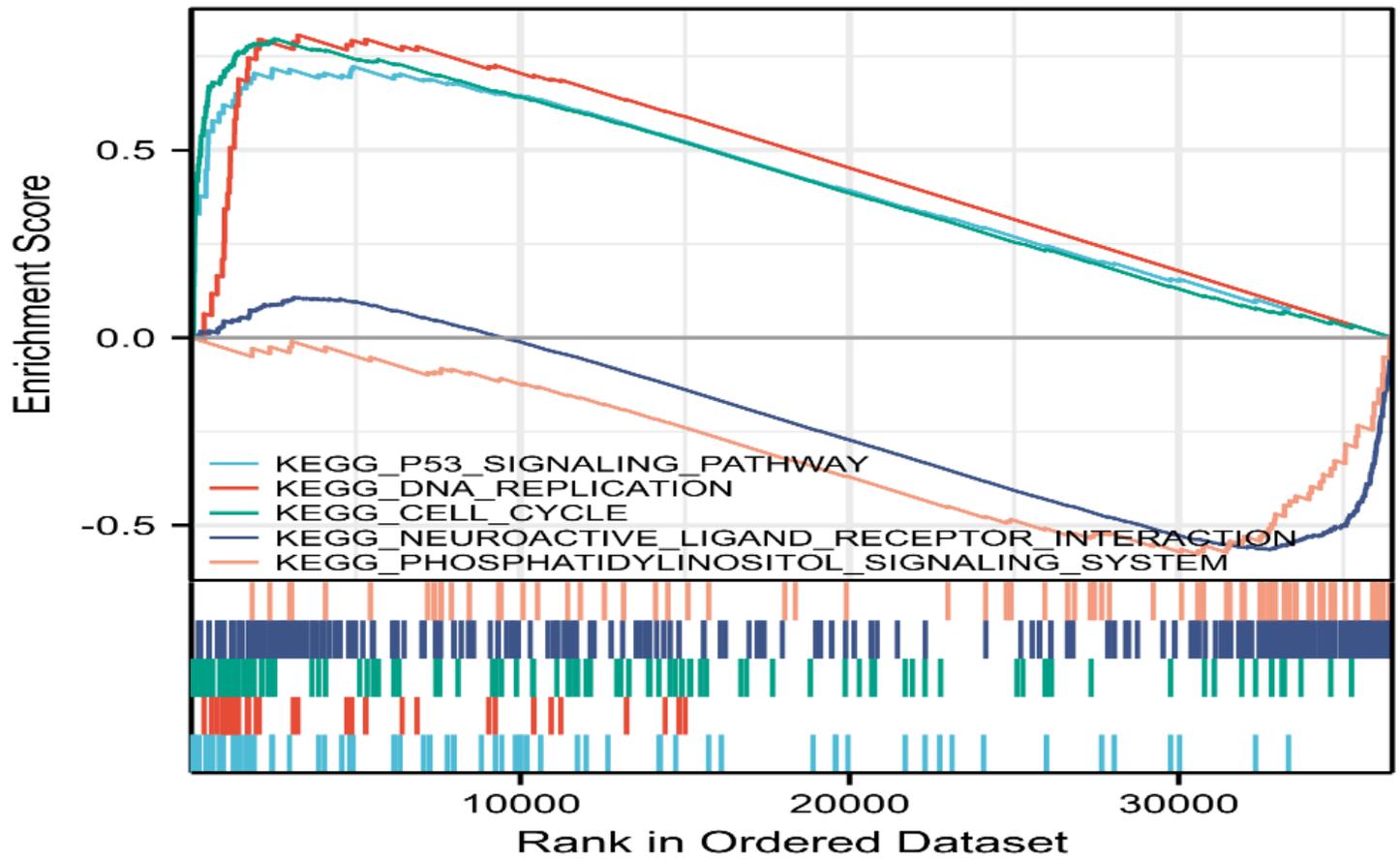


Figure 11

GSEA analysis of BIRC5 related enriched gene set.